

# Discovery of selective bioactive small molecules by targeting an RNA dynamic ensemble

Andrew C Stelzer<sup>1,4,5</sup>, Aaron T Frank<sup>2,5</sup>, Jeremy D Kratz<sup>1</sup>, Michael D Swanson<sup>3</sup>,  
Marta J Gonzalez-Hernandez<sup>3</sup>, Janghyun Lee<sup>1</sup>, Ioan Andricioaei<sup>2</sup>, David M Markovitz<sup>3</sup> &  
Hashim M Al-Hashimi<sup>1\*</sup>

**Current approaches used to identify protein-binding small molecules are not suited for identifying small molecules that can bind emerging RNA drug targets. By docking small molecules onto an RNA dynamic ensemble constructed by combining NMR spectroscopy and computational molecular dynamics, we virtually screened small molecules that target the entire structure landscape of the transactivation response element (TAR) from HIV type 1 (HIV-1). We quantitatively predict binding energies for small molecules that bind different RNA conformations and report the *de novo* discovery of six compounds that bind TAR with high affinity and inhibit its interaction with a Tat peptide *in vitro* ( $K_i$  values of 710 nM–169  $\mu$ M). One compound binds HIV-1 TAR with marked selectivity and inhibits Tat-mediated activation of the HIV-1 long terminal repeat by 81% in T-cell lines and HIV replication in an HIV-1 indicator cell line ( $IC_{50}$  ~23.1  $\mu$ M).**

RNA is growing in importance as a drug target<sup>1</sup>, but current approaches used to identify protein-targeting small molecules are ill suited for RNA. Most RNA targets lack the enzymatic activity required for conventional high-throughput screening, and their conformation-switching activity has proven difficult to assay experimentally<sup>2–4</sup>. Computational docking<sup>5</sup> can, in principle, provide the structural information needed to generally infer activity and can be used to screen uncharted regions of chemical space for novel RNA binders<sup>6</sup>. However, current protocols do not take into account the large conformational changes flexible RNA receptors undergo on binding small molecules, limiting discovery to compounds that target a narrow region of the structure landscape<sup>7–9</sup>.

There is growing evidence that small molecules trigger RNA conformational changes by binding to conformers from preexisting dynamic ensembles<sup>10–17</sup>. We have recently introduced a general approach<sup>11</sup> for visualizing RNA dynamic ensembles with the atomic resolution required for computational screening, and with extended timescales (less than milliseconds) needed to broadly sample the entire structure landscape. Here multiple sets of NMR residual dipolar coupling (RDC) data that report on the dynamics of bond vectors relative to elongated RNA helices<sup>15</sup> are used to guide selection<sup>18</sup> of conformers from a large pool generated using molecular dynamics (MD)<sup>11</sup>. By finding the minimum number of conformers that satisfy all time-averaged RDC data<sup>19</sup>, we constructed a compact ensemble that samples unique and dominant positions across the entire RNA structure landscape. This combination of experiment and theory is crucial for defining RNA ensembles, given that the conformational space that has to be sampled is vast and difficult to reduce, and that current force fields remain underdeveloped for RNA compared to proteins. Using this approach, we constructed a dynamic ensemble for the TAR (ref. 12) from HIV-1 and showed that the 20 conformers in the TAR<sup>NMR-MD</sup> ensemble include structures very similar to the structures observed when TAR is bound to seven distinct small molecules<sup>11</sup> (Fig. 1a and Supplementary Results, Supplementary Fig. 1).

Here we show that docking small molecules onto RNA dynamic ensembles generated by NMR and MD solves the problem of taking into account large degrees of RNA conformational adaptation during virtual screening. Using this approach, we successfully identify six small molecules containing unidentified RNA-binding moieties that bind TAR with high affinity and inhibit its interaction with a Tat peptide *in vitro* ( $K_i$  values ranging between 710 nM and 169  $\mu$ M). One of the compounds binds HIV-1 TAR with marked selectivity through unique interactions involving both the bulge and apical loop and specifically inhibits Tat-mediated activation of the HIV-1 long terminal repeat by 81% in T cell lines and HIV replication in an HIV-1 indicator cell line ( $IC_{50}$  ~23.1  $\mu$ M). From these studies, a new strategy emerges for selectively targeting highly flexible RNAs.

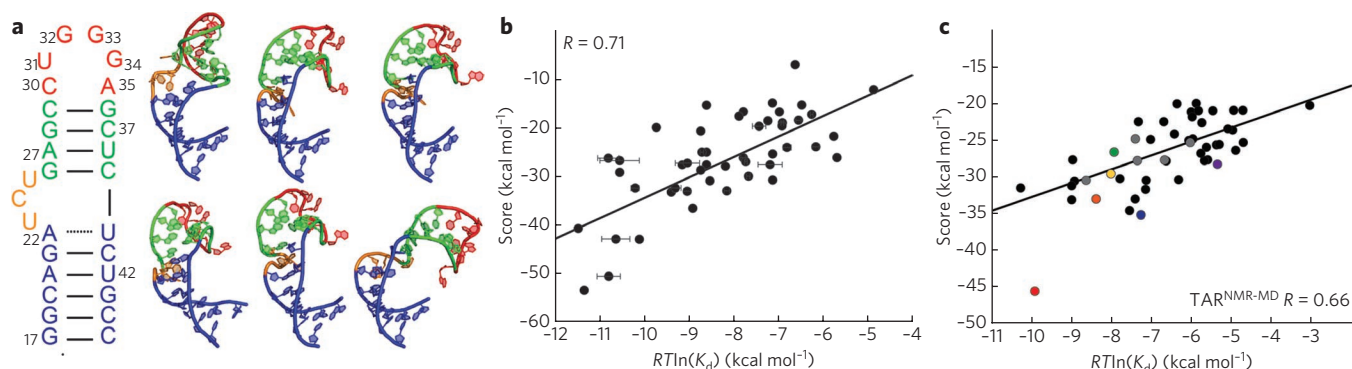
## RESULTS

### Accurate docking against known bound RNA structures

Docking small molecules onto individual conformers within NMR-MD ensembles, rather than a single static conformation, provides a straightforward but as-of-yet unrealized approach for taking into account RNA conformational adaptation during virtual screening. Such an approach inherently assumes that computational docking can be used to predict RNA–small molecule interactions with sufficient accuracy when the structure of the small molecule–bound RNA is known. We therefore benchmarked binding predictions, using the Internal Coordinate Mechanics (ICM)<sup>20</sup> docking program, for the ideal case in which the small molecule–bound RNA structure is known.

We used a diverse set of 96 small molecule–bound RNA structures, 48 of which had corresponding experimental  $K_d$  values. Structures with highly flexible small molecules ( $N_{flex} > 20$ , where  $N_{flex}$  is the number of flexible torsions in the small molecule) that pose conformational-sampling problems were excluded from analysis<sup>21</sup>. For each complex, the small molecule was removed, energy was minimized in the absence of RNA and the molecule was re-docked

<sup>1</sup>Department of Chemistry and Biophysics, University of Michigan, Ann Arbor, Michigan, USA. <sup>2</sup>Department of Chemistry, University of California Irvine, Irvine, California, USA. <sup>3</sup>Department of Internal Medicine, Division of Infectious Diseases, University of Michigan Medical Center, Ann Arbor, Michigan, USA. <sup>4</sup>Present address: Nymirum, Ann Arbor, Michigan, USA. <sup>5</sup>These authors contributed equally to this work. \*e-mail: hashimi@umich.edu



**Figure 1 | Virtually screening RNA dynamic structure ensemble.** (a) Secondary structure of wild-type HIV-1 TAR RNA, and representative conformers 3 (top), 12-15 (middle) and 18 (bottom) from the TAR<sup>NMR-MD</sup> dynamic ensemble<sup>11</sup>. (b) Correlation between experimental  $\Delta G = RT\ln(K_d)$ , with accompanying errors (s.d.) when reported, and ICM docking scores. Correlation coefficient (R) is shown for 48 RNA-small molecule complexes when small molecules were docked onto the known small molecule-bound RNA structure. (c) Correlation of  $\Delta G$  and ICM score, as in b, for 20 TAR conformers from the TAR<sup>NMR-MD</sup> ensemble.

onto the target RNA structure using fully flexible small-molecule simulations. The RNA binding pocket, defined as all heavy atoms within 5 Å of the small molecule, was held rigid. In each case, the lowest docking score obtained from a specified number of iterations sampling different small-molecule conformations and poses was recorded (Supplementary Methods). The binding energies were predicted with very good accuracy ( $R = 0.71$ ; Fig. 1b), comparable to the accuracy obtained from state-of-the-art protein docking predictions<sup>22</sup>. More than half (53%) of the predicted binding poses match the X-ray or NMR structure to within a heavy atom r.m.s. deviation cutoff of 2.5 Å (Supplementary Fig. 2a). This success rate compares well with the variability in the NMR bundle of structures, which typically results in an average r.m.s. deviation of 1.8 Å and in some cases >3 Å (ref. 23). Thus, the accuracy of docking predictions is not fundamentally limited by the scoring function or ability to sample different small molecule poses and conformations.

### Overcoming 'adaptation' problem by docking RNA ensemble

A potentially more difficult problem in RNA computational docking is that the small molecule-bound RNA structure is generally not known, and it can vary substantially from small molecule to small molecule. This uncertainty can relegate docking predictions to computational oblivion, particularly for highly flexible RNA receptors, which tend to undergo very large structural changes on binding small molecules. However, the impact of such uncertainty has never been quantified in RNA docking simulations. As an initial test, we examined how well computational docking could be used to predict the experimental binding energies for 38 TAR-binding compounds when docking against available X-ray<sup>24</sup> and NMR<sup>25</sup> structures of apo-TAR. Notably, the quality of the docking predictions deteriorated abruptly ( $R = 0.13$ ) so as to become completely uninformative and ineffective in lead-compound discovery (Supplementary Fig. 2b). Docking against a computational (TAR<sup>MD</sup>) ensemble consisting of 20 randomly chosen snapshots from an 80-ns MD simulation of apo-TAR<sup>11</sup> resulted in some improvement ( $R = 0.39$ ), but nowhere near the accuracy attainable when the bound RNA structure was known ( $R = 0.71$ ; Fig. 1b). In this approach, each small molecule was independently docked onto each of the 20 conformers, and the lowest overall score, corresponding to the dominant interaction energy, was recorded. Thus, the accuracy of docking predictions is fundamentally limited by the uncertainty in the RNA-bound structure.

We examined whether we could recover the accuracy of docking predictions by docking small molecules against the 20 conformers in the TAR<sup>NMR-MD</sup> ensemble. This NMR-informed ensemble has previously been shown to sample many of the known ligand-bound TAR

conformations<sup>11</sup>. Notably, the binding energies were now predicted with an accuracy ( $R = 0.66$ ; Fig. 1c) comparable to that attained when the bound RNA structure was known ( $R = 0.71$ ; Fig. 1b). These results reinforce the view that small molecules do not 'induce' new TAR conformations, but rather, 'capture' conformers from a pre-existing dynamic ensemble, and that TAR<sup>NMR-MD</sup> provides a good approximation for this ensemble.

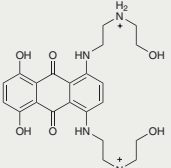
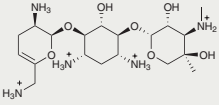
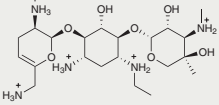
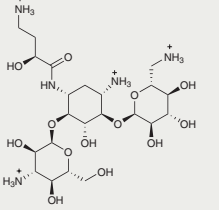
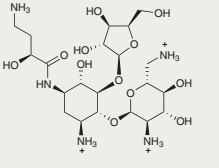
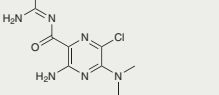
### Virtual screening TAR dynamic ensemble

The interaction between TAR and the viral transactivator protein Tat has long been targeted by researchers seeking to inhibit HIV replication, but this work has not yet resulted in clinically efficacious drugs<sup>26</sup>. We used our ensemble-targeted approach to identify TAR-targeting compounds. Each of the 20 conformers in the TAR<sup>NMR-MD</sup> ensemble was subjected to virtual screening against ~51,000 small molecules (see Methods). The top 57 commercially available hits were tested using fluorescence-based assays (Supplementary Figs. 3 and 4) that probe binding to a TAR construct containing a fluorescent probe at bulge residue U25 (ref. 27) and inhibition of the interaction between TAR and an N-terminally labeled fluorescein peptide containing the arginine-rich motif of TAR's cognate protein target Tat<sup>28</sup>. We experimentally validated six small molecules in this manner, showing that they bind TAR ( $K_d = 55$  nM–122 μM) and inhibit its interaction with Tat ( $K_i = 710$  nM–169 μM) (Table 1).

Together with spermine (Supplementary Fig. 5), the compounds were identified with a hit rate of 12%; the hit rate was as high as 50% when we focused only on water-soluble compounds that did not require DMSO in experimental assays (DMSO was not included in docking simulations). This can be compared to a hit rate of 0% in a screen of 57 randomly selected small molecules from the same libraries (Supplementary Data). The virtual screen also identified several small molecules, including two aminoglycosides, that were correctly predicted to bind TAR with much weaker affinity, despite their multiple positive charges. This was verified using fluorescence-based binding assays (Supplementary Fig. 6a).

The six small molecules include previously unidentified compounds that add to the chemical diversity of known TAR-binding small molecules (Table 1). These compounds differ from recently developed cyclic peptides that bind TAR with high affinity and specificity<sup>29,30</sup>, as well as from other TAR binders. For example, mitoxantrone (1; Table 1), a known RNA binder<sup>31</sup>, binds TAR with an affinity ( $K_d \sim 55$  nM) that, among non-neamine derivatives, is second only to the small molecule WM5 ( $K_d \sim 19$  nM)<sup>32</sup>, which was identified over two decades of research targeting TAR. The compound 5-(N,N)-dimethylamiloride (DMA; 6;  $K_d \sim 122$  μM) lacks

**Table 1 | Discovery of TAR-binding small molecules by virtually screening the TAR dynamic structure ensemble**

Structure	Compound	$K_d$ ( $\mu$ M)	$K_i$ ( $\mu$ M)
	Mitoxantrone ( <b>1</b> )	$0.055 \pm 0.021$	$0.71 \pm 0.32$
	Sisomicin ( <b>2</b> )	$0.73 \pm 0.24$	$6.4 \pm 2.7$
	Netilmicin ( <b>3</b> )	$1.35 \pm 0.35$	$14.1 \pm 8.3$
	Amikacin ( <b>4</b> )	$1.54 \pm 0.29$	$16.2 \pm 9.6$
	Butirosin A ( <b>5</b> )	$4.78 \pm 0.53$	$13.8 \pm 5.3$
	5-(N,N)-Dimethylamiloride ( <b>6</b> )	$121.85 \pm 50.65$	$169 \pm 98$

$K_d$  values were determined using fluorescence-intensity measurements (for mitoxantrone, sisomicin, netilmicin, amikacin and 5-(N,N)-dimethylamiloride) or NMR chemical-shift perturbations (for butirosin A, owing to commercial unavailability at the onset of these experiments).  $K_i$  values were determined using fluorescence-polarization measurements. Errors are s.d.

cationic groups, contains a novel RNA-binding scaffold consisting of a 5-chloropyrazin-2-amine core (Table 1) and targets a unique pocket within the TAR apical loop (see below). This is a rare example of a small molecule binding exclusively to an RNA apical loop. The molecules also include the four semisynthetic aminoglycosides sisomicin (**2**;  $K_d \sim 0.73 \mu$ M), netilmicin (**3**;  $K_d \sim 1.4 \mu$ M), amikacin (**4**;  $K_d \sim 1.5 \mu$ M) and butirosin A (**5**;  $K_d \sim 4.8 \mu$ M) (Table 1), none of which has previously been shown to bind TAR.

### Netilmicin binds TAR with high selectivity

The small molecules seem to have widely different specificities, as assayed using a competition experiment in which the  $K_d$  is re-measured after the addition of 100-fold excess tRNA<sup>33</sup> (Supplementary Fig. 7). Although we observed a substantial deterioration in the binding affinities of mitoxantrone, amikacin and sisomicin ( $K_d$  values increased by factors of 27, 7 and 3, respectively), consistent with nonspecific binding to tRNA, we observed little to no change in the  $K_d$  values for netilmicin and DMA, indicating that these compounds bind HIV-1 TAR with high specificity (Supplementary Fig. 7). Notably, a single ethyl group substantially reduces the binding specificity of sisomicin as compared to netilmicin in this assay.

As a more stringent test of specificity, we used fluorescence-based assays to measure the binding affinities of the small molecules (excluding butirosin A, which became commercially unavailable at the onset of these experiments) for a panel of three RNAs that more closely resemble the TAR hairpin. These included an HIV-2 TAR variant (HIV-2) that differs from HIV-1 TAR by deletion of a single bulge residue, insertion of a G-U base pair and swapping of a G-C base-pair in the upper stem; the prokaryotic ribosomal A-site hairpin (A-site); and the HIV-1 rev response element hairpin (RRE). Both the A-site and RRE are binding sites for a broad range of aminoglycosides<sup>34,35</sup> (Fig. 2a).

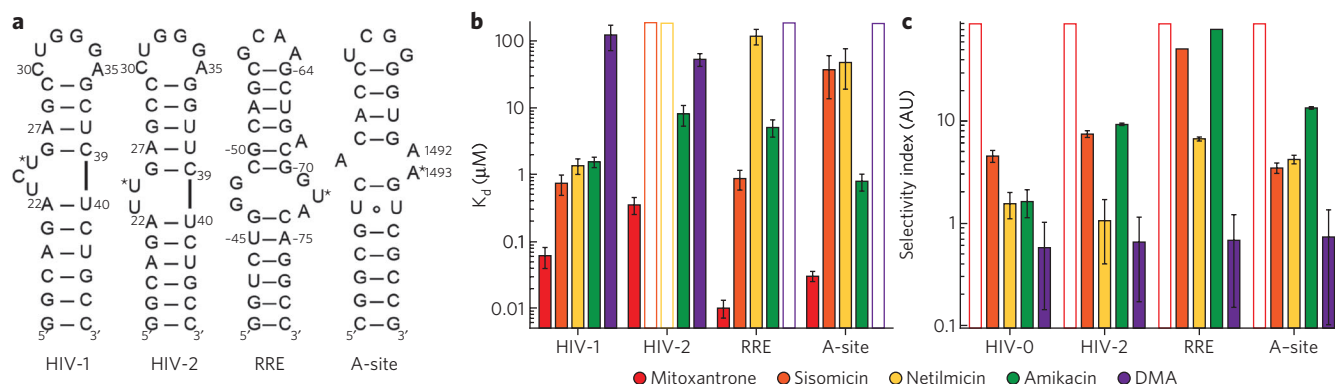
The binding assays (Supplementary Fig. 8a) yielded specificity profiles for the various compounds that mirrored those observed with tRNA (Supplementary Fig. 7a). Netilmicin showed the highest selectivity. It bound the closely related HIV-2 TAR with negligible affinity, and A-site and RRE RNA with one-thirty-fifth and one-eighty-sixth the affinity of HIV-1 TAR, respectively (Fig. 2b,c and Supplementary Fig. 8a). As an even more stringent test, we measured the binding affinity of netilmicin to a TAR mutant that had a deletion of a single cytosine bulge residue. This mutant bound netilmicin with one-sixteenth the affinity of wild-type TAR (Supplementary Fig. 8a). Once again, sisomicin showed markedly less specificity than netilmicin; it bound RRE with an affinity comparable to that for HIV-1 TAR (Fig. 2b,c). As expected, DMA, which binds to the TAR apical loop (see below), showed strong selectivity against A-site and RRE but not against HIV-2 TAR (Fig. 2b,c). All the other small molecules bound at least one other RNA with an affinity comparable to that for HIV-1 TAR (Supplementary Fig. 8a). We confirmed these specificity trends with competition assays analogous to those described for tRNA, using the three RNA constructs as competitors (Supplementary Fig. 7a).

### Testing predicted binding modes using NMR

We tested the docking-predicted TAR-small molecule binding modes with site-specific resolution using NMR chemical shift mapping experiments. Many of the small molecules were predicted to bind conformers within a contiguous region (conformers 12–15) of the TAR<sup>NMR-MD</sup> structure landscape (Fig. 3a) characterized by near-coaxial alignment of the helices (interhelical bend angle  $<12^\circ$ ), as observed for TAR when bound to Tat mimics. Accordingly, all of the small molecules induced chemical shift perturbations characteristic of coaxial stacking of TAR helices as observed with Tat peptides and divalent ions<sup>36</sup> (for example, U23 and C24, Fig. 3b and Supplementary Fig. 9). Notably, netilmicin, which showed the highest TAR binding specificity *in vitro*, was also predicted to bind conformers within the TAR ensemble with the highest specificity among aminoglycosides, with one conformer (18) accounting for 66% of the TAR population.

All six small molecules were predicted to bind TAR using distinct modes and to contact various combinations of residues in the bulge, upper stem and apical loop (Fig. 3c) that form crucial interactions with Tat, providing a structural basis for inhibiting the TAR-Tat interaction. Accordingly, the small molecules induce distinct chemical shift perturbations, particularly for residues predicted to be within the binding pocket (Fig. 3b,c and Supplementary Fig. 9). Substantial perturbations ( $>0.1$  p.p.m.) were observed for 87% of TAR sites predicted to be within 5 Å of the small molecule (Fig. 3c, red spheres). Perturbations outside the 5-Å cutoff (Fig. 3c, green spheres) typically corresponded to nearby flexible residues, which probably change conformation on binding the small molecule. Even detailed aspects of the predicted binding modes are supported, in certain cases, by the NMR data (Fig. 3b,c and Supplementary Methods), including unique stacking of mitoxantrone on G26, distinct binding modes for netilmicin and sisomicin mediated by contacts involving netilmicin's unique ethyl group, and binding of DMA to a unique pocket within the TAR apical loop. Unlike other



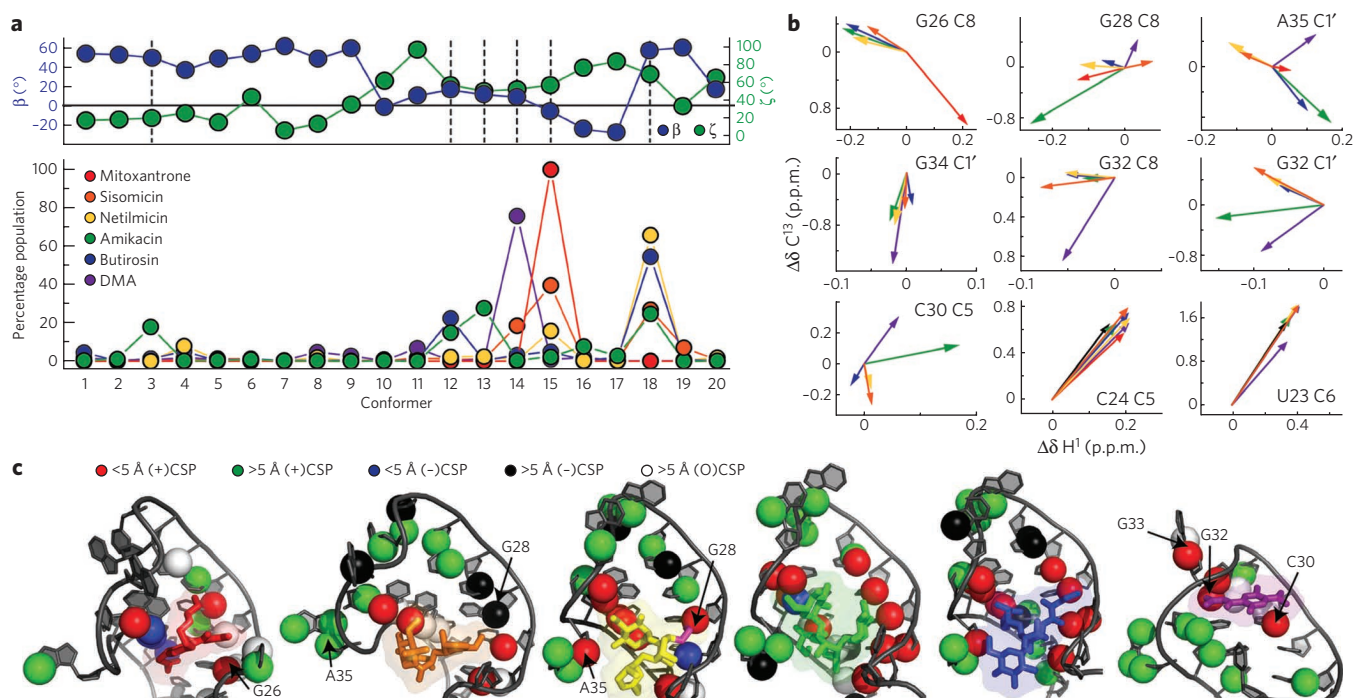


**Figure 2 | Analysis of small-molecule binding specificity.** (a) Secondary structure of three related RNA hairpins (HIV-2 TAR, RRE and the prokaryotic A-site) used in binding-specificity analysis. Asterisks indicate sites of 2'-fluorescein and 2-aminopurine base substitutions. (b)  $K_d$  values for each small molecule binding to four RNA constructs, determined using fluorescence-intensity measurements. Open bars indicate a  $K_d$  could not be determined because the small molecule bound the RNA too weakly to saturate binding. Error bars denote s.d. (c) Selectivity index (ratio of the measured  $K_d$  of a small molecule binding HIV-1 TAR in the presence of competitor RNA over the  $K_d$  in the absence of competitor) for each small molecule. AU, arbitrary units. Open bars indicate that a  $K_d$  could not be determined reliably owing to strong competition and incomplete HIV-1 TAR saturation. Error bars denote s.d. The apparently higher specificity of DMA can be attributed to binding to the apical loop and therefore reduced susceptibility to competition from compounds that primarily target the bulge.

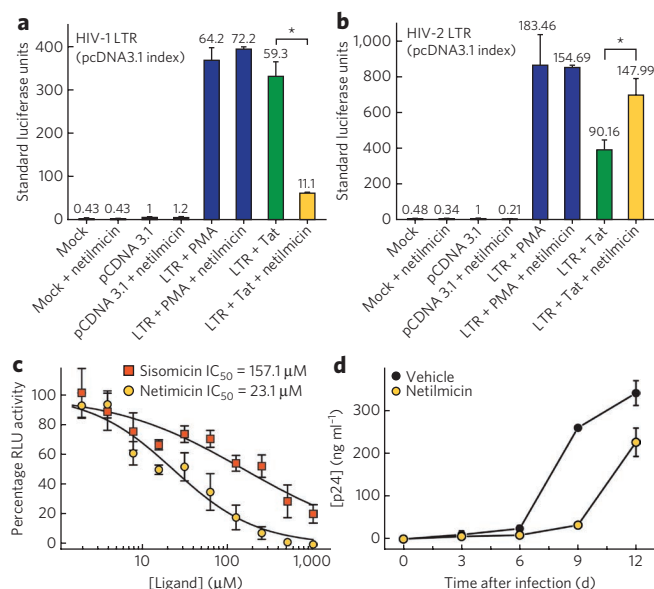
aminoglycosides and their conjugated derivatives that have been shown to bind the TAR bulge and upper and lower stems<sup>37–39</sup>, the aminoglycosides identified here, including netilmicin, interact with the apical loop in addition to the bulge and upper stem, while DMA provides a rare example of a small molecule that exclusively targets an RNA apical loop.

### Netilmicin inhibits Tat activation of HIV-1 LTR

Of the five compounds tested (excluding butirosin A, which became commercially unavailable at the onset of these experiments), netilmicin, which binds TAR with the highest specificity *in vitro* (Fig. 2b), inhibited Tat-mediated activation of the HIV-1 promoter by ~81% compared to the control (Fig. 4a), as assayed in live human



**Figure 3 | NMR site-specific characterization of TAR-small molecule binding modes.** (a) Docking-predicted preferences for the binding of small molecules to distinct conformers in the TAR<sup>NMR-MD</sup> ensemble. Top, the interhelical bend ( $\beta$ ) and twist ( $\zeta$ ) angles are plotted for each TAR conformer. Bottom, the predicted fractional population of each small molecule bound to each of the 20 TAR conformers, assuming the ICM computed energy. (b) Representative examples of chemical shift perturbation vectors, colored according to the small molecule showing perturbations (as in a), in resonance positions from free TAR to >97% small molecule-bound TAR.  $\text{Mg}^{2+}$  perturbations are shown in black. For clarity, NMR peaks are not shown and are provided in Supplementary Figure 9. (c) Mapping the NMR chemical shift perturbation data onto the docking-predicted TAR-small molecule structures. TAR residues are colored according to whether they are predicted to be within or outside a 5-Å distance cutoff from any atom in the small molecule, and whether they show substantial (>0.1 p.p.m.) chemical shift perturbations (+CSP) or not (-CSP). (O) denotes that perturbations could not be assigned owing to spectral overlap. The ethyl substituent of netilmicin is colored purple.



**Figure 4 | Netilmicin specifically inhibits Tat-mediated transactivation and HIV-1 replication.** (a) Jurkat T cells were treated with 100 μM netilmicin or vehicle, for 24 h and then transfected with the HIV-1 LTR-luciferase construct (LTR) and the HIV-1 Tat construct (Tat) or with empty vector (pcDNA3.1). ‘Mock’ refers to the untransfected control. After transfection, netilmicin or vehicle was added to the media, and luciferase activity was measured 24 h later. As a control, cells were transfected with only LTR and, 4 h before the 24-h completion point, were treated with 10 ng ml<sup>-1</sup> PMA, with or without netilmicin. \*P < 0.02. (b) Experiments as in a, using HIV-2 LTR with HIV-2 Tat. Error bars in a and b are s.e.m. from three independent transfections. (c) Netilmicin inhibits HIV-1 replication in the HIV-1 indicator cell line TZM-bl, which expresses luciferase upon HIV-1 infection. Error bars denote s.d. and were obtained from duplicate measurements. For comparison, results with sisomicin are also shown. RLU, relative light units. (d) Netilmicin inhibits HIV-1 replication in the T-cell line Hut78. Cells were infected with the HIV-1 isolate NL4-3. HIV-1 replication was assessed at 3-d intervals by p24 ELISA. Error bars denote s.d. and were calculated from duplicate measurements.

T cells using a luciferase reporter construct transfected into Jurkat T cells. The other four compounds did not show activity in this assay, probably owing to their much weaker binding specificities, though we cannot rule out other effects, such as differences between full-length Tat used in transfection assays and Tat peptides used in the *in vitro* displacement assay. As a control, we repeated measurements of netilmicin activity upon addition of phorbol 12-myristate 13-acetate (PMA), which activates the HIV-1 long terminal repeat (LTR) in a Tat-independent manner. If netilmicin does indeed block Tat-mediated activation of the HIV-1 promoter through its interaction with TAR, then no inhibition should be observed upon PMA-mediated activation. Indeed, netilmicin did not inhibit PMA-mediated stimulation of the HIV-1 promoter, thus ruling out off-target effects (Fig. 4a).

To further assess the specificity of netilmicin, we measured its inhibitory activity toward an HIV-2 TAR promoter containing the same HIV-2 TAR sequence used in the binding assays (Fig. 2b). According to *in vitro* binding data (Fig. 2b), netilmicin binds HIV-2 TAR with negligible affinity. If the mode of action of netilmicin involves binding to TAR, we would predict it would be far less effective at inhibiting the stimulation of the HIV-2 promoter. Indeed, netilmicin did not inhibit stimulation of the HIV-2 promoter; rather, slightly greater activity was observed (Fig. 4b). Moreover, no inhibition was observed when we used HIV-1 Tat to stimulate the HIV-2 transcriptional promoter, indicating that netilmicin does not bind and inhibit HIV-1 Tat, but rather, affects

Tat-mediated transactivation through its interaction with HIV-1 TAR (Supplementary Fig. 10). Thus, netilmicin specifically inhibits its activation of the HIV-1 promoter but not the closely related HIV-2 promoter in cellular assays.

### Netilmicin inhibits HIV-1 replication

Notably, netilmicin inhibited not only Tat activation, but also HIV-1 replication, as assayed using an HIV-1 indicator cell line, TZM-bl, and the HIV-1 NL4-3 isolate, which contains the same HIV-1 TAR sequence used in the *in vitro* studies. Addition of netilmicin to cells before and during infection resulted in substantially less HIV replication, yielding an IC<sub>50</sub> value (~23.1 μM; Fig. 4c) strikingly similar to the value measured *in vitro* in the Tat displacement assay (K<sub>i</sub> = 14.1 μM; Table 1). The similarity of these inhibition constants further supports the idea that netilmicin inhibits HIV replication by targeting TAR. This IC<sub>50</sub> compares favorably with EC<sub>50</sub> values (ranging between 0.7 μM and 30 μM) measured using the same HIV-1 NL4-3 strain for the most potent aminoglycoside derivatives that have been designed as Tat mimetics, including, for example, the aminoglycoside-arginine conjugate of neomycin (NeoR6; EC<sub>50</sub> ~0.7 μM)<sup>40</sup>.

We further corroborated the above results in an *in vivo* assay by infecting the HUT-78 T-cell line with HIV-1 NL4-3 in the presence of 100 μM netilmicin. Every 3 d, we assessed HIV-1 replication by measuring the amount of p24 antigen in the culture supernatants. Smaller amounts of p24 were observed for samples treated with netilmicin compared to vehicle alone, with the largest difference observed on day 9 (Fig. 4d). Inhibition by netilmicin toxicity rather than TAR binding was ruled out, as trypan blue staining for cellular viability showed little difference between the vehicle- and netilmicin-treated cells (data not shown). Finally, our *in vitro* and cellular (gene reporter) assays showed that, despite its chemical similarity to netilmicin, sisomicin was far less effective at inhibiting Tat-mediated activation of the HIV-1 promoter, most probably because it binds TAR with substantially less specificity. If netilmicin inhibits HIV-1 replication by inhibiting the TAR-Tat interaction, we would expect sisomicin to be a far less potent inhibitor. Indeed, the IC<sub>50</sub> value (~157.1 μM) measured for sisomicin in the same HIV replication assay is about seven-fold higher than that measured for netilmicin (Fig. 4c). This difference is substantial, considering that a two-fold increase in the IC<sub>50</sub> is typically observed with resistant viruses. These data provide additional support for the notion that netilmicin inhibits HIV replication by selectively inhibiting the TAR-Tat interaction, unlike NeoR and other aminoglycoside-arginine conjugates, for which data suggest a different mode of inhibition involving the blocking of viral entry<sup>40</sup>.

### DISCUSSION

Netilmicin is, to our knowledge, the first experimentally validated RNA-targeting compound with *in vivo* activity to be identified using a virtual screen. It has exquisite binding selectivity for HIV-1 TAR, and this seems to be an important determinant of its activity. This high specificity is achieved in part via a single ethyl substituent on a key cationic amine group (Table 1). Comparison of the binding modes of netilmicin and sisomicin reveals that this modification alone substantially reduces netilmicin's binding affinity for tRNA (Supplementary Fig. 7a) and RRE (Fig. 2b,c), without affecting its binding affinity for HIV-1 TAR. The alkyl group may stereochemically block access to the cationic group in the more rigid and less adaptable small molecule-binding pockets of tRNA and RRE, but not in the more flexible and malleable HIV-1 TAR. Consistent with this notion, reducing TAR's flexibility by deleting a single cytosine bulge residue that was not observed to make direct contacts with netilmicin, reduced netilmicin's binding affinity to one-sixteenth of its affinity for wild-type TAR (Supplementary Fig. 8a). Previous studies have shown that the flexible TAR, compared with the less

malleable A-site, is better able to accommodate conformationally restrained small molecules<sup>41</sup>. Stereochemical crowding of key cationic groups on small molecules may well prove to be a general strategy for enhancing selectivity toward highly flexible RNAs.

The method developed here for targeting highly flexible RNA receptors can also be implemented to target other highly flexible targets, including intrinsically unfolded proteins implied in neurodegenerative diseases for which traditional structure- and assay-based approaches have thus far failed, and for which NMR-informed computational dynamic ensembles are beginning to emerge<sup>42</sup>. Although five of the six newly identified TAR binders reported here have previously been shown to bind other RNAs, the compound library used in our virtual screen has been optimized for protein high-throughput screening and is not enriched with compounds that have favorable nucleic acid-binding properties. Our virtual screen can be scaled up to include millions of compounds that have advantageous nucleic acid-binding and drug-like characteristics. This will provide a much-needed route for efficiently screening new regions of chemical space in the search for novel RNA-targeting lead compounds.

## METHODS

**Virtual screening of TAR dynamic ensemble.** Virtual screening simulations were performed with ICM (Molsoft)<sup>20</sup> using 20 TAR<sup>NMR-MD</sup> conformers and ~51,000 small molecules. TAR binding pockets were defined using the ICM PocketFinder module, and the protonation states of the small molecules were computed over pH 5.4–9.4 using the Major Microspecies module in the Calculator Plugin (ChemAxon). The small-molecule library (totaling 51,226 compounds) used in the virtual screening consisted of 49,166 compounds obtained from the Center for Chemical Genomics at University of Michigan and 2,060 compounds from our in-house library. Small molecules in the in-house library were identified from published reports of verified RNA-binding small molecules and drawn using ChemDraw (CambridgeSoft). Both libraries were saved in sdf file format. We started the virtual screening simulations by docking small molecules with  $N_{\text{flex}} < 20$  using a thoroughness of 1, followed by a second round of screening small molecules with the top ~10% of scores using a thoroughness of 10. The top 57 small molecules (58 including spermine) were subjected to experimental validation. The 57 small molecules were obtained from Maybridge, Chembridge, LKT Labs and Sigma, and with the exception of sisomicin (purity ≥80%), all were guaranteed to be ≥95% pure.

**Fluorescence-based binding assay.** The binding assay used a TAR construct labeled with 2-aminopurine at bulge residue U25. An alternative TAR construct labeled with fluorescein at the same residue was used to measure binding of small molecules whose absorbance spectra overlapped with the fluorescence of 2-aminopurine. Both RNAs were purchased from Integrated DNA Technologies. RNA was annealed by heating at 95 °C for 5 min followed by dilution (100 nM) into working buffer (10 mM phosphate, 20 mM NaCl, 0.1 mM EDTA pH 6.8) and cooling on ice for 2 h. Samples were pre-equilibrated for 5 min after addition of small molecule. Fluorescence intensity was measured using a Fluoromax-2 fluorimeter at an excitation wavelength of 320 nm and emission wavelength of 390 nm for 2-aminopurine, and an excitation wavelength of 485 nm and emission wavelength of 520 nm for fluorescein. Fluorescence-intensity measurements were recorded in triplicate and normalized to those of unbound TAR.

**Fluorescence-based TAR-Tat displacement assay.** The fluorescence polarization-based displacement assay used an N-terminally fluorescein-labeled Tat peptide (N-AAARKRRRR-C, Genscript) and an *in vitro*-synthesized elongated TAR. The elongated TAR was used to increase the dynamic range of fluorescence-polarization measurements. The TAR (60 nM) was incubated with varying concentrations of small molecules for 10 min, followed by another 10-min incubation with the addition of fluorescein-labeled Tat peptide (10 nM). The fluorescence-polarization buffer consisted of 50 mM Tris, 100 mM NaCl and 0.01% (v/v) nonidet-P40 at pH 7.4. The UV absorption spectrum was recorded for each small molecule tested, to ensure there was no spectral overlap with fluorescein. Fluorescence polarization was measured in triplicate using 384-well plates read with a PHERAstar Plus plate reader (BMG Labtech) with a 485-nm excitation wavelength and 520-nm detection wavelength for the optic module.  $IC_{50}$  and  $K_i$  values were calculated using the Prism software (GraphPad Software).  $K_i$  values and corresponding errors reported in Table 1 are within the 95% confidence interval.

**NMR experiments.** All NMR experiments were performed at 25 °C with an Avance Bruker 600-MHz spectrometer equipped with a 5-mm triple-resonance cryogenic probe. NMR buffer consisted of 15 mM sodium phosphate, 0.1 mM EDTA, 25 mM NaCl and 10% D<sub>2</sub>O at pH ~6.4. NMR spectra were processed and analyzed with NMRPipe<sup>43</sup> and SPARKY 3 (<http://www.cgl.ucsf.edu/home/sparky/>).

**Cellular assays.** In transfection assays, cells were pretreated with small molecule, or with vehicle or water (as controls), 24 h before transfection. Data shown in Figure 4a,b and Supplementary Figure 10 were normalized to *Renilla* luciferase activity and represent the average of three independent transfections. Student's *t* test, comparing Tat and netilmicin treatment to Tat and vehicle treatment, was used to obtain the *P* values in Figure 4a,b and Supplementary Figure 10. The HIV-1 indicator cell line TZM-bl, which expresses luciferase upon HIV-1 infection, was plated in 96-well plates and treated with netilmicin, or water as a control, 24 h before infection. The cells were then infected with HIV-1 NL4-3. Forty-eight hours after infection, luciferase activity was quantified as relative light units. Values were normalized and an  $IC_{50}$  value computed using nonlinear regression. The T-cell line Hut78 was infected with the HIV-1 isolate NL4-3, and HIV-1 replication was assessed using p24 ELISA. At 3-d intervals, half of the medium was harvested and replaced with uninfected Hut78 cells, with or without 100 μM netilmicin.

Received 11 December 2010; accepted 18 April 2011;  
published online 26 June 2011

## References

- Cooper, T.A., Wan, L. & Dreyfuss, G. RNA and disease. *Cell* **136**, 777–793 (2009).
- Parsons, J. *et al.* Conformational inhibition of the hepatitis C virus internal ribosome entry site RNA. *Nat. Chem. Biol.* **5**, 823–825 (2009).
- Blount, K.F. & Breaker, R.R. Riboswitches as antibacterial drug targets. *Nat. Biotechnol.* **24**, 1558–1564 (2006).
- Thomas, J.R. & Hergenrother, P.J. Targeting RNA with small molecules. *Chem. Rev.* **108**, 1171–1224 (2008).
- Kuntz, I.D. Structure-based strategies for drug design and discovery. *Science* **257**, 1078–1082 (1992).
- Filikov, A.V. *et al.* Identification of ligands for RNA targets via structure-based virtual screening: HIV-1 TAR. *J. Comput. Aided Mol. Des.* **14**, 593–610 (2000).
- Hermann, T. Rational ligand design for RNA: the role of static structure and conformational flexibility in target recognition. *Biochimie* **84**, 869–875 (2002).
- Cruz, J.A. & Westhof, E. The dynamic landscapes of RNA architecture. *Cell* **136**, 604–609 (2009).
- Fulle, S. & Gohlke, H. Molecular recognition of RNA: challenges for modelling interactions and plasticity. *J. Mol. Recognit.* **23**, 220–231 (2010).
- Zhang, Q., Sun, X., Watt, E.D. & Al-Hashimi, H.M. Resolving the motional modes that code for RNA adaptation. *Science* **311**, 653–656 (2006).
- Frank, A.T., Stelzer, A.C., Al-Hashimi, H.M. & Andricioaei, I. Constructing RNA dynamical ensembles by combining MD and motionally decoupled NMR RDCs: new insights into RNA dynamics and adaptive ligand recognition. *Nucleic Acids Res.* **37**, 3670–3679 (2009).
- Puglisi, J.D., Tan, R., Calnan, B.J., Frankel, A.D. & Williamson, J.R. Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science* **257**, 76–80 (1992).
- Williamson, J.R. Induced fit in RNA-protein recognition. *Nat. Struct. Biol.* **7**, 834–837 (2000).
- Leulliot, N. & Varani, G. Current topics in RNA-protein recognition: Control of specificity and biological function through induced fit and conformational capture. *Biochemistry* **40**, 7947–7956 (2001).
- Zhang, Q., Stelzer, A.C., Fisher, C.K. & Al-Hashimi, H.M. Visualizing spatially correlated dynamics that directs RNA conformational transitions. *Nature* **450**, 1263–1267 (2007).
- Latham, M.P., Zimmermann, G.R. & Pardi, A. NMR chemical exchange as a probe for ligand-binding kinetics in a theophylline-binding RNA aptamer. *J. Am. Chem. Soc.* **131**, 5052–5053 (2009).
- Vaiana, A.C. & Sanbonmatsu, K.Y. Stochastic gating and drug-ribosome interactions. *J. Mol. Biol.* **386**, 648–661 (2009).
- Chen, Y., Campbell, S.L. & Dokholyan, N.V. Deciphering protein dynamics from NMR data using explicit structure sampling and selection. *Biophys. J.* **93**, 2300–2306 (2007).
- Clare, G.M. & Schwieters, C.D. Amplitudes of protein backbone dynamics and correlated motions in a small alpha/beta protein: correspondence of dipolar coupling and heteronuclear relaxation measurements. *Biochemistry* **43**, 10678–10691 (2004).
- Abagyan, R., Totrov, M. & Kuznetsov, D. ICM—a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem.* **15**, 488–506 (1994).
- Lang, P.T. *et al.* DOCK 6: combining techniques to model RNA-small molecule complexes. *RNA* **15**, 1219–1230 (2009).
- Cheng, T., Li, X., Li, Y., Liu, Z. & Wang, R. Comparative assessment of scoring functions on a diverse test set. *J. Chem. Inf. Model.* **49**, 1079–1093 (2009).
- Guilbert, C. & James, T.L. Docking to RNA via root-mean-square-deviation-driven energy minimization with flexible ligands and flexible targets. *J. Chem. Inf. Model.* **48**, 1257–1268 (2008).



24. Ippolito, J.A. & Steitz, T.A. A 1.3-Å resolution crystal structure of the HIV-1 trans-activation response region RNA stem reveals a metal ion-dependent bulge conformation. *Proc. Natl. Acad. Sci. USA* **95**, 9819–9824 (1998).
25. Aboul-ela, F., Karn, J. & Varani, G. Structure of HIV-1 TAR RNA in the absence of ligands reveals a novel conformation of the trinucleotide bulge. *Nucleic Acids Res.* **24**, 3974–3981 (1996).
26. Yang, M. Discoveries of Tat-TAR interaction inhibitors for HIV-1. *Curr. Drug Targets Infect. Disord.* **5**, 433–444 (2005).
27. Bradrick, T.D. & Marino, J.P. Ligand-induced changes in 2-aminopurine fluorescence as a probe for small molecule binding to HIV-1 TAR RNA. *RNA* **10**, 1459–1468 (2004).
28. Matsumoto, C., Hamasaki, K., Mihara, H. & Ueno, A. A high-throughput screening utilizing intramolecular fluorescence resonance energy transfer for the discovery of the molecules that bind HIV-1 TAR RNA specifically. *Bioorg. Med. Chem. Lett.* **10**, 1857–1861 (2000).
29. Davidson, A., Patora-Komisarska, K., Robinson, J.A. & Varani, G. Essential structural requirements for specific recognition of HIV TAR RNA by peptide mimetics of Tat protein. *Nucleic Acids Res.* **39**, 248–256 (2011).
30. Davidson, A. *et al.* Simultaneous recognition of HIV-1 TAR RNA bulge and loop sequences by cyclic peptide mimics of Tat protein. *Proc. Natl. Acad. Sci. USA* **106**, 11931–11936 (2009).
31. White, R.J. & Durr, F.E. Development of mitoxantrone. *Invest. New Drugs* **3**, 85–93 (1985).
32. Parolin, C. *et al.* New anti-human immunodeficiency virus type 1 6-aminoquinolones: mechanism of action. *Antimicrob. Agents Chemother.* **47**, 889–896 (2003).
33. Blount, K.F., Tor, Y., Hamasaki, K. & Ueno, A. Using pyrene-labeled HIV-1 TAR to measure RNA-small molecule binding aminoglycoside antibiotics, neamine and its derivatives as potent inhibitors for the RNA-protein interactions derived from HIV-1 activators. *Nucleic Acids Res.* **31**, 5490–5500 (2003).
34. DeJong, E.S., Chang, C.E., Gilson, M.K. & Marino, J.P. Proflavine acts as a Rev inhibitor by targeting the high-affinity Rev binding site of the Rev responsive element of HIV-1. *Biochemistry* **42**, 8035–8046 (2003).
35. Kaul, M., Barbieri, C.M. & Pilch, D.S. Fluorescence-based approach for detecting and characterizing anti biotic-induced conformational changes in ribosomal RNA: Comparing aminoglycoside binding to prokaryotic and eukaryotic ribosomal RNA sequences. *J. Am. Chem. Soc.* **126**, 3447–3453 (2004).
36. Stelzer, A.C., Kratz, J.D., Zhang, Q. & Al-Hashimi, H.M. RNA dynamics by design: biasing ensembles towards the ligand-bound state. *Angew. Chem. Int. Ed. Engl.* **49**, 5731–5733 (2010).
37. Lapidot, A., Berchanski, A. & Borkow, G. Insight into the mechanisms of aminoglycoside derivatives interaction with HIV-1 entry steps and viral gene transcription. *FEBS J.* **275**, 5236–5257 (2008).
38. Lapidot, A., Vijayabaskar, V., Litovchick, A., Yu, J.G. & James, T.L. Structure-activity relationships of amino glyco side-arginine conjugates that bind HIV-1 RNAs as determined by fluorescence and NMR spectroscopy. *FEBS Lett.* **577**, 415–421 (2004).
39. Faber, C., Sticht, H., Schweimer, K. & Rosch, P. Structural rearrangements of HIV-1 Tat-responsive RNA upon binding of neomycin B. *J. Biol. Chem.* **275**, 20660–20666 (2000).
40. Cabrera, C. *et al.* Anti-HIV activity of a novel aminoglycoside-arginine conjugate. *Antiviral Res.* **53**, 1–8 (2002).
41. Blount, K.F., Zhao, F., Hermann, T. & Tor, Y. Conformational constraint as a means for understanding RNA-aminoglycoside specificity. *J. Am. Chem. Soc.* **127**, 9818–9829 (2005).
42. Boehr, D.D., Nussinov, R. & Wright, P.E. The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **5**, 789–796 (2009).
43. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on unix pipes. *J. Biomol. NMR* **6**, 277–293 (1995).

## Acknowledgments

We thank A.V. Kurochkin for NMR expertise, and we thank the Michigan Economic Development Cooperation and the Michigan Technology Tri-Corridor for support of the purchase of a 600-MHz spectrometer. This work was supported by the US National Institutes of Health (R01 AI066975-01 and R01 CA144043), the US National Science Foundation (NSF Career Award CHE-0918817) and an NSF Graduate Research Fellowship for A.C.S. and A.T.F.

## Author contributions

H.M.A.-H. and A.C.S. conceived the docking approach; A.T.F., I.A., A.C.S. and H.M.A.-H. developed the approach for constructing RNA dynamic structure ensembles; A.C.S., with input from A.T.F. and J.D.K., performed the docking simulations; A.C.S., J.D.K. and J.L. performed the *in vitro* fluorescence assays and NMR experiments; M.D.S., M.J.G.-H. and D.M.M. carried out the transfection and viral-replication assays; H.M.A.-H., A.C.S., A.T.F., J.D.K., I.A. and D.M.M. wrote the paper.

## Competing financial interests

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at <http://www.nature.com/naturechemicalbiology/>.

## Additional information

Supplementary information and chemical compound information is available online at <http://www.nature.com/naturechemicalbiology/>. Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>. Correspondence and requests for materials should be addressed to H.M.A.-H.